AD-A049 479 STANFORD UNIV CALIF F/G 12/1

A QUEUEING SYSTEM SUBJECT TO BREAKDOWN AND HAVING NON-STATIONAR--ETC(U)

NOV 77 A W SHOGAN N00014-75-C-0561

UNCLASSIFIED TR-188 NL

1 OF 1

ADA049 479

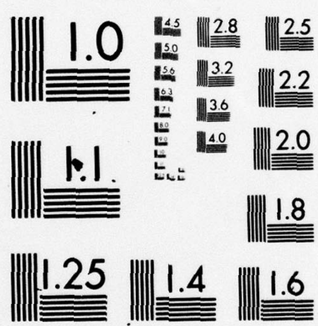
END

DATE

FILMED

3 -78

DDC


1.0
4.5
5.0
5.6
6.3
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

A QUEUEING SYSTEM SUBJECT TO BREAKDOWN AND HAVING NON-STATIONARY POISSON ARRIVALS

by

Andrew W. Shogan

TECHNICAL REPORT NO. 188

November 7, 1977

SUPPORTED BY THE ARMY AND NAVY
UNDER CONTRACT N00014-75-C-0561 (NR-042-002)
WITH THE OFFICE OF NAVAL RESEARCH

Gerald J. Lieberman, Project Director

DEPARTMENT OF OPERATIONS RESEARCH
AND
DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

A QUEUEING SYSTEM SUBJECT TO BREAKDOWN AND HAVING NONSTATIONARY POISSON ARRIVALS

Andrew W. Shogan
University of California, Berkeley

In many queueing systems (e.g., a computer facility), the server is subject to breakdown. If the breakdowns are unpredictable in nature and the server is not replaced or repaired until a breakdown occurs, then the facility will be unable to provide uninterrupted service to its customers. In such a case, it is important to understand how the breakdowns will effect the system's level of performance.

Avi-Itzhak and Naor [1] and Gaver [3] obtained the steady state quantities L and W for an M/G/1 queue whose unreliable server alternates between operational periods of exponential duration and failed periods of arbitrary random duration. Although general with respect to the distributions of the service and repair times, the results have the disadvantage of requiring a stationary Poisson arrival process.

Yechiali and Naor [7] and Fond and Ross [2] both considered variations of a single server exponential queueing model in which the arrival and service rates alternate between the pairs (λ_1, μ_1) and (λ_2, μ_2), with the length of time the system operates with pair (λ_i, μ_i) being exponentially distributed with rate $c\alpha_i$, i = 1, 2. In [7], the steady state quantity L was derived for a system with infinite queue capacity while, in [2], the steady-state proportion of customers lost was obtained for a system in which any arrival finding the server busy is lost. The

special case $\mu_2 = 0$ results in models of queueing systems subject to breakdown and having nonstationary Poisson arrivals; however, all probability distributions are restricted to be exponential.

The model developed here is an important one that cannot be obtained as a special case of the models in the literature. In particular, consider a single server queueing system with the following characteristics:

(a) The system alternates between two states: the operational state and the failed state (also referred to as the repair state).

(b) When operational, the system functions as an $M/E_k/1$ queue; that is, customers arrive according to a Poisson process with rate λ, and service is according to an Erlang distribution with mean $1/\mu$ and shape parameter k.

(c) If service to a customer is interrupted by a breakdown, resumption takes place as soon as the repair period ends with no loss of service involved.

(d) Although no service takes place during the repair period, customers continue to arrive according to a Poisson process, but now having a rate of λ_1 instead of λ.

(e) The duration of operating periods is exponential with mean $1/c\alpha$ and the duration of repair periods is Erlang with mean $1/c\beta$ and shape parameter m.

This model is not a special case of [7] because it permits both the service and repair times to have Erlang distributions rather than restricting them to be exponential. Furthermore, the model is not a special case of

[1] and [3] because the Poisson arrival process is nonstationary when $\lambda_1 \neq \lambda$. A nonstationary arrival process is useful in many practical situations where the customers are aware when the server is inoperable. In such cases, one expects to find $\lambda_1 < \lambda$ or even $\lambda_1 = 0$.

The constant c in assumption (e) controls how rapidly the system occilates between the operational and failed states. Holding α and β constant while increasing c has the effect of keeping the steady state probabilities of being in the operational and failed states constant while at the same time increasing the frequency with which the system changes states.

The remainder of this paper is organized as follows: Section 1 analyzes the queueing system described by assumptions (a) - (e); because Little's formula is not valid for this model unless $\lambda_1 = \lambda$, the steady-state quantities L and W must be derived separately. Some special and extreme cases of the general model are considered in Section 2. Section 3 not only shows that both L and W are decreasing and convex functions of c but also investigates the behavior of the system as $c \to \infty$. The paper concludes in Section 4 with an application to a production-storage system.

1. STEADY STATES RESULTS

Preliminary Analysis

Throughout this section, as well as Sections 2 and 4, the constant c will be assumed to equal 1. This eliminates the need to carry c along in all the derivations when it is only relevant to the analysis in Section 3.

The "method of phases" (cf. [4, p. 168]) provides a convenient means of obtaining the steady state results. It is well-known that an Erlang random variable with mean $1/\mu$ and shape parameter k is equivalent to the sum of k independent, exponentially distributed random variables each having the same mean $1/k\mu$. Hereafter, both the Erlang service times and Erlang repair times of the model will be viewed as consisting of a series of identical and independent, exponentially distributed phases.

The system can now be analyzed as a continuous time Markov process with states $\{(i,j) | i = 0,1, \ldots, m$ and $j = 0, 1, 2, \ldots\}$ where $i = 0$ denotes the system is operational, $1 \leq i \leq m$ denotes the number of phases remaining in the repair process until the system becomes operational, and j denotes the number of service phases in the system (the sum of the number of phases remaining for the customer in service and k times the number of customers in the queue). The transition probabilities are stationary and satisfy the Kolmogorov differential equations. Furthermore, the steading state probabilities $\{p_{ij}\}$ exist, are independent of the initial state, and satisfy the following balance equations:

$$(\lambda_1+m\beta)\ p_{m0} = \alpha\ p_{00} \qquad (j = 0) \qquad (1a)$$

$$(\lambda_1+m\beta)\ p_{i0} = m\beta\ p_{i+1,0} \quad (1 \leq i \leq m-1) \qquad (j = 0) \qquad (1b)$$

$$(\lambda+\alpha)\ p_{00} = m\beta\ p_{10} + k\mu\ p_{01} \qquad (j = 0) \qquad (1c)$$

$$(\lambda_1+m\beta)\ p_{mj} = \alpha\ p_{0j} + \lambda_1\ p_{m,j-k} \qquad (j > 0) \qquad (1d)$$

$$(\lambda_1+m\beta)\ p_{ij} = m\beta\ p_{i+1,j} + \lambda_1\ p_{i,j-k} \quad (1 \leq i \leq m-1) \quad (j > 0) \qquad (1e)$$

$$(\lambda+\alpha+k\mu)\ p_{0j} = m\beta\ p_{1j} + \lambda\ p_{0,j-k} + k\mu p_{0,j+1} \qquad (j > 0) \qquad (1f)$$

where a negative subscript in (1d) - (1f) indicates the term is zero. Figure 1 contains the portion of the Markov chain's state transition diagram corresponding to states (i,j) with $0 \leq i \leq m$ and $j \geq k$. It is clear from the figure that equations (1) can be interpreted as requiring the mean transition rates into and out of a state to be equal at steady state.

Let $p_w = \sum_{j=0}^{\infty} p_{0j}$ and $p_f = \sum_{i=1}^{m} \sum_{j=0}^{\infty} p_{ij}$; that is, p_w and p_f are the steady state probabilities of the system being operational and failed, respectively. On considering the underlying two state (operational and failed) stochastic process, it is immediate that

$$p_w = \beta/(\alpha+\beta)\ ,$$

$$p_f = \alpha/(\alpha+\beta)\ .$$

Let the average arrival rate and average service rate in steady state be denoted by $\hat{\lambda} = \lambda p_w + \lambda_1 p_f$ and $\hat{\mu} = \mu p_w$; furthermore, let $r = \hat{\lambda}/\hat{\mu}$. It will be demonstrated shortly that, as is often the case, $\hat{\lambda} < \hat{\mu}$ is a condition for steady state. Note that each of the quantities p_w, p_f, $\hat{\lambda}$, $\hat{\mu}$, and r would be independent of c even if the temporary assumption $c = 1$ were dropped.

The Generating Function

Generating-function techniques must be used to further analyze the model as there is no way of solving (1) in a recursive manner to obtain closed-form expressions for the $\{p_{ij}\}$. Define the generating functions

$$G_i(z) = \sum_{j=0}^{\infty} p_{ij} z^j \qquad |z| \leq 1, \quad i = 0,1,2,\ldots,m$$

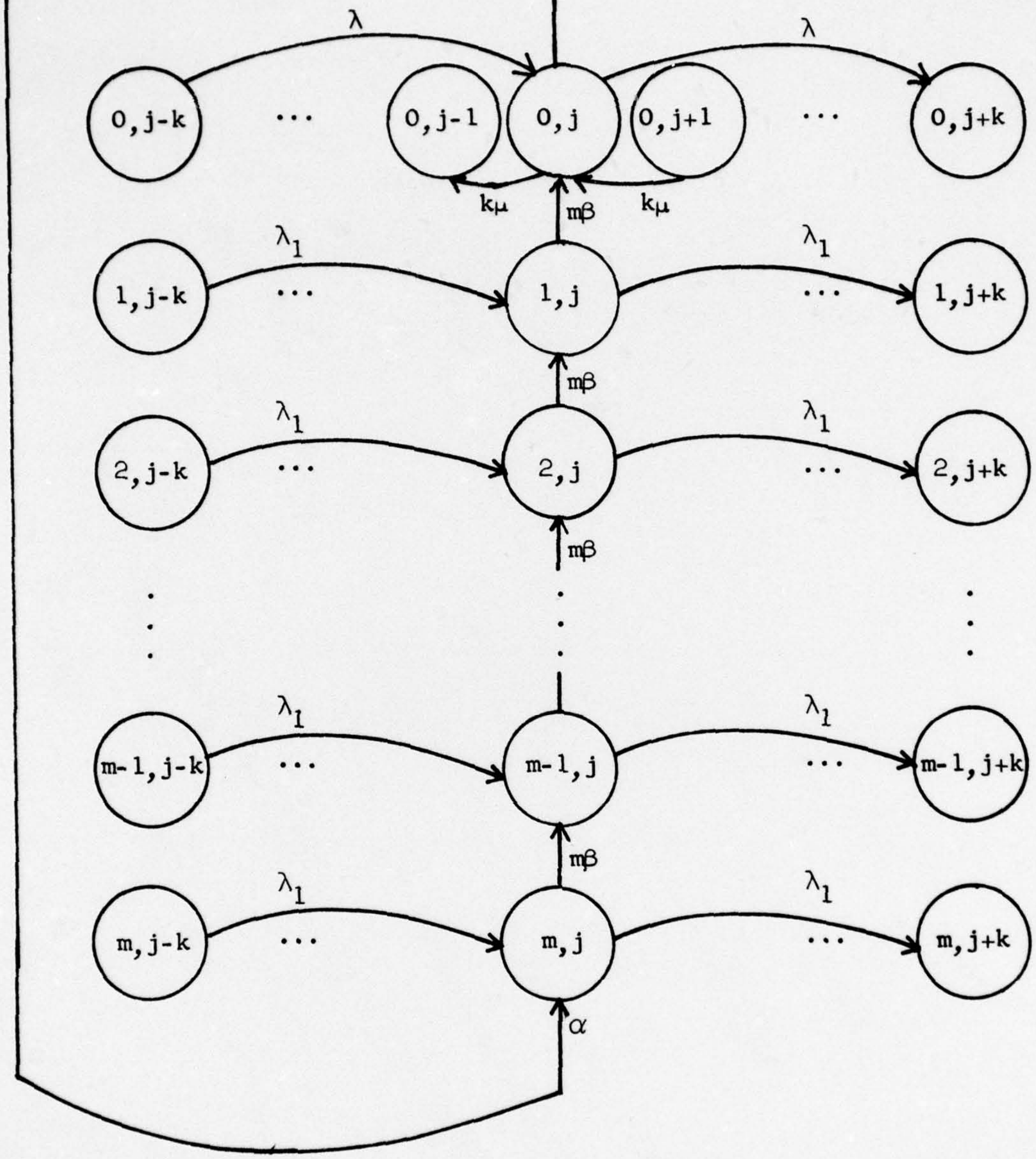
$$G(z) = \sum_{i=0}^{m} G_i(z) \qquad |z| \leq 1 \,. \tag{2}$$

Multiplying each equation of the sets $\{(1a), (1d)\}$, $\{(1b), (1e)\}$, and $\{(1c), (1f)\}$ by z^j and summing over all j yields, respectively,

$$G_m(z) = [\alpha/(\lambda_1 + m\beta - \lambda_1 z^k)]\, G_0(z) \,, \tag{3}$$

$$G_i(z) = [m\beta/(\lambda_1 + m\beta - \lambda_1 z^k)]\, G_{i+1}(z) \,, \qquad (1 \leq i \leq m-1) \,, \tag{4}$$

$$G_0(z) = [m\beta z\, G_1(z) - k\mu\, p_{00}(1-z)]/[(\lambda + \alpha + k\mu)z - \lambda z^{k+1} - k\mu] \,. \tag{5}$$



λ
λ
0,j-k
0,j-1
0,j
0,j+1
0,j+k
kμ
mβ
kμ
λ1
λ1
1,j-k
1,j
1,j+k
mβ
λ1
λ1
2,j-k
2,j
2,j+k
mβ
λ1
λ1
m-1,j-k
m-1,j
m-1,j+k
mβ
λ1
λ1
m,j-k
m,j
m,j+k
α

Figure 1

Equations (3) and (4) can be used recursively to express $G_1(z)$ in terms of $G_0(z)$ as

$$G_1(z) = (\alpha/m\beta)[m\beta/(\lambda_1+m\beta-\lambda_1 z^k)]^m \, G_0(z) \tag{6}$$

and (5) can be rearranged as

$$G_1(z) = (m\beta z)^{-1} \{[(\lambda+\alpha+k\mu)z - \lambda z^{k+1} - k\mu] \, G_0(z) + k\mu p_{00}(1-z)\}. \tag{7}$$

Equating the expressions for $G_1(z)$ in (6) and (7) and solving for $G_0(z)$ gives

$$G_0(z) = \{k\mu \; p_{00}(1-z) \; [f(z)]^m\}/D(z), \tag{8}$$

where

$$f(z) = 1 + (\lambda_1/m\beta) \; (1-z^k) \, , \tag{9}$$

and the denominator is

$$D(z) = \alpha z + [k\mu + \lambda z^{k+1} - (\lambda+\alpha+k\mu)z] \; [f(z)]^m \; . \tag{10}$$

Using (3) and (4) recursively results in

$$G_i(z) = \{(\alpha/m\beta) \; k\mu \; p_{00}(1-z) \; [f(z)]^{i-1}\}/D(z) \quad (1 \le i \le m). \tag{11}$$

$G(z)$ can now be calculated from (2), (8), and (11) as

$$G(z) = p_{00}[N(z)/D(z)] \quad , \tag{12}$$

where

$$N(z) = k\mu(1-z) \{[f(z)]^m + (\alpha/m\beta) \sum_{i=0}^{m-1} [f(z)]^i\} \quad . \tag{13}$$

Using $G(1) = 1$ and computing $\lim_{z \to 1} G(z)$ by applying L'Hôpital's rule to (12) yields, after algebraic simplification,

$$p_{00} = (1-r)\ p_w \quad . \tag{14}$$

Substituting (14) into (12) results in a final expression for $G(z)$, that is

$$G(z) = (1-r)\ [\beta/(\alpha+\beta)]\ [N(z)/D(z)] \ . \tag{15}$$

Since $p_{00} > 0$, (14) also verifies the previously mentioned condition for steady state, $\hat{\lambda} < \hat{\mu}$.

The busy fraction $\rho \equiv 1 - \sum_{i=0}^{m} p_{i0}$ equals $1 - G(0)$ and can be evaluated from (15), (9), (10), and (13). Provided $\lambda_1 > 0$,

$$\rho = 1 - (1-r)\ [\beta/(\alpha+\beta)]\ \{1 + (\alpha/\lambda_1)\ [1 - (m\beta/(\lambda_1+m\beta))^m]\} \quad .$$

In general, then, $\rho \neq (\hat{\lambda}/\hat{\mu})$; however, if $\lambda_1 = 0$, $\rho = (\hat{\lambda}/\hat{\mu})$ does hold.

Recursions for the $\{p_{ij}\}$

Unfortunately, no simple relationship exists relating the $\{p_{ij}\}$ to $p_{00}, p_{10}, \ldots, p_{m0}$. However, the $\{p_{ij}\}$ can be computed efficiently from (14) and the set of recursive equations (for $j = 0, 1, 2, \ldots$)

$$p_{mj} = (\lambda_1+m\beta)^{-1} (\alpha p_{0j} + \lambda_1 p_{m,j-k}) \tag{16a}$$

$$p_{ij} = (\lambda_1+m\beta)^{-1} (m\beta\, p_{i+1,j} + \lambda_1\, p_{i,j-k}), \quad (i=m-1,m-2,\ldots,1) \tag{16b}$$

$$p_{0,j+1} = (k\mu)^{-1}(\lambda \textstyle\sum_{n=j+1-k}^{j} p_{0n} + \alpha \sum_{n=0}^{j} p_{0n} - m\beta \sum_{n=0}^{j} p_{1n}) \tag{16c}$$

where a term is zero if it has a negative subscript and the lower limit of summation in the first term of (16c) is reset to 0 if it is negative. Equations (16a) and (16b) are obviously equivalent to (1d) and (1e) while (16c) follows from (1f) and a simple inductive argument. Of course, the steady-state probability of having n customers in the system is given by $1-\rho$ for $n = 0$ and, for $n > 0$, by $\sum_{i=0}^{m} \sum_{j=(n-1)k+1}^{nk} p_{ij}$.

Computation of L

The computation of L and L_q, the steady-state average number of customers in the system and in the queue, respectively, require some preliminary results. Let L^P and L_q^P be steady-state notation for the average number of customer service phases in the system and in the queue, respectively. Furthermore, let L_s^P denote the average number of service

phases remaining for the customer (if any) in service. The following relationships clearly hold:

$$L_q^P = kL_q \; , \tag{17}$$

$$L^P = L_q^P + L_s^P \tag{18}$$

$$L = L_q + \rho \; . \tag{19}$$

Substituting (17) into (18) and solving for L_q results in

$$L_q = (1/k)\,(L^P - L_s^P) \; . \tag{20}$$

Relationships (19) and (20) then yield

$$L = (1/k)\,(L^P - L_s^P) + \rho \; . \tag{21}$$

The problem, then, is to compute L^P and L_s^P.

Now $L^P = G'(1)$; however, evaluating $G'(1)$ is not easy. Expressing $G(z) = p_{00}\,N(z)/D(z)$ and using L'Hôpital's rule twice gives

$$G'(1) = p_{00}[N''(1)\cdot D'(1) - N'(1)\cdot D''(1)]/2[D'(1)]^2 \; . \tag{22}$$

The algebraic manipulations required by (22) are straightforward but quite long. Because they would occupy several pages, the computations are omitted; however, they result in

$$L^P = \frac{r}{1-r}\left[\frac{k+1}{2} + \frac{m+1}{2m} \cdot \frac{k\alpha\lambda_1(\mu-\lambda+\lambda_1)}{\hat{\lambda}(\alpha+\beta)^2}\right] . \tag{23}$$

Obtaining L_s^P requires the development of another generating function. In the $M/E_k/1$ queue not subject to breakdown, given that a customer is in service in steady-state, the number of phases remaining until his service is complete is equally likely to be 1, 2, ..., or k (cf. [4, p. 169]). However, this is not the case when the queue is subject to breakdown. Define the generating functions

$$H_i(y) = p_{i0} + \sum_{n=1}^{k} \left(\sum_{j=0}^{\infty} p_{i,jk+n}\right) y^n \quad , \quad |y| \leq 1,\ i = 0,1,\ldots,m$$

$$H(y) = \sum_{i=0}^{m} H_i(y) \quad , \quad |y| \leq 1 . \tag{24}$$

Clearly, $L_s^P = H'(1)$. As demonstrated in the Appendix, the lengthy derivation of $H(y)$ results in

$$H(y) = (1-\rho) + (\rho-r)y^k + r\{[y(1-y^k)]/[k(1-y)]\} . \tag{25}$$

Application of L'Hôpital's rule twice to (25) yields

$$L_s^P = H'(1) = (\rho-r)k + r[(k+1)/2] . \tag{26}$$

Finally, combining (21), (23), and (26) results in

$$L = \frac{r}{1-r}\left[(1-r) + \frac{k+1}{2k} \cdot r + \frac{m+1}{2m} \cdot \frac{\alpha\lambda_1(\mu+\lambda_1-\lambda)}{\hat{\lambda}(\alpha+\beta)^2}\right] . \tag{27}$$

Computation of W

Because Little's formula is not valid for this model unless $\lambda_1 = \lambda$, L cannot be used to immediately obtain W, the steady-state expected value of the time a customer spends both in the queue and in service. Instead, let $W = T_1 + T_2 + T_3$ where the T_i $(i = 1, 2, 3)$ are the expected values in steady state for

1. The time that elapses from the customer's arrival until the queue first becomes operational; if the queue is operational when the customer arrives, this time equals zero.

2. The time that elapses from when the queue first becomes operational until everyone present when the customer arrived has been served; if the arriving customer finds the system empty, this time equals zero.

3. The time that elapses from when the customer's service begins until it ends.

Each of the T_i $(i = 1, 2, 3)$ will now be calculated.

Applying L'Hôpital's rule to (11) yields $G_i(1) = \alpha/[m(\alpha+\beta)]$ for $i = 1, 2, \ldots, m$ so that, given the system is failed, the number of exponential repair phases remaining until it becomes operational is equally distributed over $1, 2, \ldots, m$. Hence, given the system is failed, the expected number of repair phases remaining is $[(m+1)/2]$. Since each repair phase has an expected length of $(1/m\beta)$,

$$T_1 = [(m+1)/2m\beta]\ [\alpha/(\alpha+\beta)]\ . \tag{28}$$

The times T_2 and T_3 can both be divided into two parts: time during which the system is operational (T_i^w; $i = 2, 3$) and time during which the system is failed (T_i^f; $i = 2,3$). Clearly, $T_2^w = L^P/k\mu$ and $T_3^w = 1/\mu$. During the expected operational times T_i^w ($i = 2, 3$), the expected number of failures is αT_i^w. Since the expected repair time for each failure is $1/\beta$, $T_i^f = (\alpha/\beta) T_i^w$ for $i = 2, 3$. Hence,

$$T_2 = [(\alpha+\beta)/\beta]\,[L^P/k\mu] \quad , \tag{29}$$

$$T_3 = [(\alpha+\beta)/\beta\mu] \ . \tag{30}$$

Combining (23), (28), (29) and (30) results in

$$W = (\hat{\lambda})^{-1}\left(\frac{r}{1-r}\right)\left\{(1-r) + \frac{k+1}{2k}\cdot r + \frac{m+1}{2m}\cdot\frac{\alpha}{\alpha+\beta}\right.$$

$$\left.\cdot\left[\frac{\mu}{\alpha+\beta} + \left(\frac{\lambda_1(\mu+\lambda_1-\lambda) - \hat{\lambda}\mu}{\beta\mu}\right)\right]\right\}. \tag{31}$$

Note from (27) and (31) that $L = \hat{\lambda}W$ does <u>not</u> hold unless $\lambda_1 = \lambda$.

2. SPECIAL AND EXTREME CASES

<u>Case A</u>. As $\beta \to \infty$, the repair periods have shorter and shorter durations, and, in the limit, repair is instantaneous. Intuitively, then, as $\beta \to \infty$, the model developed in Section 1 approaches the $M/E_k/1$ queue <u>not</u> subject to breakdown and having constant arrival rate λ and service rate μ. That this is in fact the case can be shown from (15), (27), and (31); as $\beta \to \infty$, $G(\cdot)$, L, and W all approach the corresponding quantities for the $M/E_k/1$ queue.

<u>Case B</u>. If $k = 1$ and/or $m = 1$, exponential service and/or repair times result. When both $k = 1$ and $m = 1$, expression (27) for L reduces to a special case $(\mu_2 = 0)$ of expression (33) of Yechiali and Naor [7, p. 729].

<u>Case C</u>. Constant service and/or repair times can be analyzed by letting $k \to \infty$ and/or $m \to \infty$. It is clear from (27) and (31) that L and W are decreasing and convex functions of both k and m. Their limiting values are obtained by replacing $[(k+1)/2k]$ and/or $[(m+1)/2m]$ by $1/2$ in (27) and (31).

<u>Case D</u>. If the Poisson arrival rate is stationary $(\lambda_1 = \lambda)$, expressions (27) and (31) for L and W reduce to special cases (Erlang service and repair) of relationships (24) and (26) of Avi-Itzhak and Naor [1, p. 309].

Case E. In some practical situations, no customers enter the system when it is failed, either by their own choice or because of restrictions by the system. During the repair process, then, customers neither enter nor leave the system. Thus, it is intuitive that not only L but also the steady-state probabilities of having n customers in the system ($n = 0, 1, 2, \ldots$) are equal to those for the $M/E_k/1$ queue not subject to breakdown and having constant arrival rate λ and service rate μ. That this is in fact the case can be seen by setting $\lambda_1 = 0$. Then $r = \lambda/\mu$ and, from (15) and (27), both $G(z)$ and L simplify to the corresponding quantities for the $M/E_k/1$ queue. Of course, as (31) with $\lambda_1 = 0$ indicates, W is greater than in the $M/E_k/1$ case.

3. BEHAVIOR OF L AND W AS A FUNCTION OF c

In order to investigate how the system behaves as a function of c, the assumption of Sections 1, 2, and 4 that $c = 1$ is now dropped. Recall that c controls how rapidly the system occilates between the operational and failed states. Varying c while holding α and β constant does not change p_w and p_f, the steady-state probabilities of the system being operational and failed, respectively. However, as c increases, the system fluctuates more rapidly between the operational and failed states, or, equivalently, the mean time the system stays in each state approaches 0.

Expressions for L and W as a function of c can be obtained by replacing α and β by $c\alpha$ and $c\beta$ everywhere in (27) and (31). From the expressions that result, it is easily shown that

$$L'(c) = -ac^{-2}$$

$$W'(c) = -bc^{-2}$$

where a and b are positive constants involving m, λ, λ_1, μ, α, and β. Hence, both L and W are decreasing and convex functions of c.

To interpret this result qualitatively, consider two equally reliable systems (i.e., identical p_w's) also having identical k, m, λ, λ_1, and μ; however, suppose one system has infrequent failures but long repair times (a low c) and the other undergoes frequent but quickly repaired failues (a high c). If the objective is to minimize L and W, then the latter system should be chosen.

On a more quantitative level, the result supports a general conjecture of Ross [6] that, in a single server infinite capacity queueing model, the "more stationary" the Poisson arrival process is, then the smaller the average customer delay. As in [2] and [6], this conjecture has been verified in a special case. To see this, the behavior of the system as $c \to \infty$ will be investigated. It is easy to show from (27) and (31) that as $c \to \infty$, both L and W approach the corresponding quantities for an $M/E_k/1$ queue <u>not</u> subject to breakdown and having constant arrival rate $\hat{\lambda}$ and service rate $\hat{\mu}$. Thus, as $c \to \infty$, the system

becomes more stationary in the sense that it behaves more and more like the $M/E_k/1$ queue with parameters $\hat{\lambda}$ and $\hat{\mu}$. Also, note that since L and W are decreasing in c, the smallest values they can ever achieve are the corresponding values for the $M/E_k/1$ queue with parameters $\hat{\lambda}$ and $\hat{\mu}$.

4. APPLICATION TO A PRODUCTION-STORAGE SYSTEM

By regarding the server as a production process turning out items one at a time and each customer as a unit demand for the product, the results of Section 1 can be used to analyze a production-storage process subject to breakdown and having the following additional characteristics:

(a) Unsatisfied demand is always backlogged.

(b) Items not needed immediately to satisfy backlogged demand are stored for future use up to a level of S, the finite capacity of the storage facility.

(c) When the storage facility is filled, no production takes place.

In such a production-storage model, two quantities of interest are I, the average inventory level where a negative value indicates that the items are backlogged rather than in storage, and R, the fraction of time that demand can be met without backlogging. For example, it may be desired to choose S so that I and/or R are greater than some specified design parameters. Meyer, Rothkopf, and Eldredge [5] consider four

production-storage models differing from the one just described in that production and demand both occur at constant rates, unsatisfied demand is lost, and the durations of operating and failed periods have distributions corresponding to the four possible combinations of constant and exponential.

A state (i,j) of the queueing system translates into a state of the production-storage system as follows: $0 \leq j \leq kS$ is equivalent to $kS-j$ phases in storage and $j \geq kS$ is equivalent to $j-kS$ phases backlogged. Of course, i has the same interpretation in both models. Given this one-to-one correspondence between states in the two models, expression (27) and recursions (16) can be used to compute

$$I = S-L$$

$$R = \sum_{i=0}^{m} \sum_{j=0}^{k(S-1)} p_{ij} \qquad (S > 0)$$

where a negative value for I indicates a backlog. As a numerical example, consider a system with $\alpha = 1$, $\beta = 3$, $m = 5$, $\lambda = 8$, $\lambda_1 = 4$, $\mu = 16$,, $k = 4$ and suppose it is desired to choose S so that both $I \geq 3$ and $R \geq 0.95$. Computed with the aid of (27) and (16), the following table shows I and R as a function of S:

S	0	1	2	3	4	5	6	7	8	9	10
I	-1.45	-0.45	0.55	1.55	2.55	3.55	4.55	5.55	6.55	7.55	8.55
R	0.000	0.367	0.639	0.797	0.886	0.936	0.965	0.980	0.989	0.994	0.997

From the table, it is clear that $S \geq 6$ must hold.

APPENDIX: DERIVATION OF $H(y)$

To obtain $H(y)$ multiply each equation of the sets {(1a), (1d)}, {(1b), (1e)}, and {(1c), (1f)} by y^n where

$$n = \begin{cases} 0 & \text{if } j = 0 \\ k & \text{if } j = k, 2k, 3k, \ldots \\ j \text{ modulo } k & \text{otherwise} . \end{cases}$$

Summing over all j then yields, respectively,

$$H_m(y) = (m\beta)^{-1} [\alpha H_0(y) - \lambda_1 p_{m0}(1-y^k)] \tag{32}$$

$$H_i(y) = H_{i+1}(y) - (m\beta)^{-1} \lambda_1 p_{i0}(1-y^k) \qquad (1 \le i \le m-1) \tag{33}$$

$$H_0(y) = [\alpha y - k\mu(1-y)]^{-1} \{m\beta y H_1(y) - k\mu p_{00}(1-y) - y(1-y^k) \cdot [\lambda p_{00} + k\mu(H_0'(0) - p_{01})]\} . \tag{34}$$

Equations (32) and (33) can be used recursively to express $H_1(y)$ in terms of $H_0(y)$ as

$$H_1(y) = (m\beta)^{-1} [\alpha H_0(y) - \lambda_1(1-y^k) \textstyle\sum_{i=1}^{m} p_{i0}] \tag{35}$$

and (34) can be rearranged as

$$H_1(y) = (m\beta y)^{-1} \{[\alpha y - k\mu(1-y)]\, H_0(y) + k\mu\, p_{00}(1-y)$$

$$+ y(1-y^k)\, [\lambda p_{00} + k\mu(H_0'(0) - p_{01})]\} \; . \qquad (36)$$

Equating the expressions for $H_1(y)$ in (35) and (36) solving for $H_0(y)$ results in

$$H_0(y) = p_{00} + \{[y(1-y^k)]/[k\mu(1-y)]\}$$

$$\cdot \{\lambda_1 \textstyle\sum_{i=1}^{m} p_{i0} + \lambda p_{00} + k\mu[H_0'(0) - p_{01}]\} \; . \qquad (37)$$

Using (32) and (33) recursively yields

$$H_i(y) = (m\beta)^{-1}\, [\alpha H_0(y) - \lambda_1(1-y^k) \textstyle\sum_{n=i}^{m} p_{n0}] \; , \quad (1 \leq i \leq m) \quad (38)$$

$H(y)$ can now be obtained from (24), (37), and (38) as

$$H(y) = [(\alpha+\beta)/\beta] \left\{ p_{00} + \frac{y(1-y^k)}{k\mu(1-y)} [\lambda p_{00} + \lambda_1 \textstyle\sum_{i=1}^{m} p_{i0} + k\mu(H_0'(0) - p_{01})] \right\}$$

$$- (\lambda_1/m\beta)\,(1-y^k) \textstyle\sum_{i=1}^{m} i p_{i0} \qquad . \qquad (39)$$

Using $H(1) = 1$ and computing $\lim_{y \to 1} H(y)$ by applying L'Hôpital's rule to (39) results, after algebraic simplification, in

$$k\mu[H_0'(0) - p_{01}] = [(\beta/\alpha+\beta) - p_{00}]\mu - \lambda_1 \sum_{i=1}^{m} p_{i0} - \lambda\, p_{00} \ . \tag{40}$$

Substituting (40) into (39) gives

$$H(y) = [(\alpha+\beta)/\beta]p_{00} + r\{[y(1-y^k)]/[k(1-y)]\}$$

$$- (\lambda_1/m\beta)\ (1-y^k)\ \sum_{i=1}^{m} ip_{i0} \ . \tag{41}$$

Equations (1a) and (1b) can be rewritten as

$$m\beta\ p_{m0} = \alpha p_{00} - \lambda_1\ p_{m0}$$

$$m\beta\ p_{i0} = m\beta\ p_{i+1,0} - \lambda_1\ p_{i0} \qquad (1 \leq i \leq m-1)$$

which is equivalent (by induction) to

$$p_{i0} = (m\beta)^{-1}\ (\alpha p_{00} - \lambda_1 \sum_{n=i}^{m} p_{n0}) \qquad (i = m, m-1, \ldots, 1). \tag{42}$$

Using (42) to evaluate $\sum_{i=0}^{m} p_{i0}$ yields the relationship

$$1-\rho = [(\alpha+\beta)/\beta]p_{00} - (\lambda_1/m\beta) \sum_{i=1}^{m} ip_{i0} \ . \tag{43}$$

Because of (43) and (14), (41) simplifies to (25), the expression for $H(y)$ given in Section 1.

REFERENCES

1. B. Avi-Itzhak and P. Naor, "Some Queueing Problems with the Service Stations Subject to Breakdown," Opns. Res. 11, 303-320 (1963).

2. S. Fond and S. Ross, "A Heterogeneous Arrival and Service Queueing Loss Model," Technical Report ORC 77-12, Operations Research Center, University of California, Berkeley, CA 94720, May 1977.

3. D.P. Gaver, Jr., "A Waiting Line with Interrupted Service, Including Priorities," J. Roy. Stat. Soc. B24, 73-90 (1962).

4. D. Gross and C.M. Harris, Fundamentals of Queueing Theory, John Wiley and Sons, New York, 1974.

5. R.R. Meyer, M.H. Rothkopf, and G.G. Eldredge, "The Reliability of a Production-Storage System," Computer Sciences Technical Report No. 246, University of Wisconsin, Madison, WI 53706, March 1975.

6. S. Ross, "Average Delay in Queues with Nonstationary Poisson Arrivals," Technical Report ORC 77-13, Operations Research Center, University of California, Berkeley, CA 94720, May 1977.

7. U. Yechiali and P. Naor, "Queueing Problems with Heterogeneous Arrivals and Service," Opns. Res. 19, 722-734 (1971).

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 14 TR-188	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 A QUEUEING SYSTEM SUBJECT TO BREAKDOWN AND HAVING NON-STATIONARY POISSON ARRIVALS.		5. TYPE OF REPORT & PERIOD COVERED 9 Technical Report,
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 Andrew W. Shogan		8. CONTRACT OR GRANT NUMBER(s) 15 N00014-75-C-0561
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Operations Research and Dept. of Statistics -- Stanford University Stanford, California 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-042-002)
11. CONTROLLING OFFICE NAME AND ADDRESS Statistics and Probability Program Code 436 Office of Naval Research Arlington, Virginia 22217		12. REPORT DATE 11 7 Nov 77 13. NUMBER OF PAGES 23 12 28p.
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION IS UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

QUEUEING SYSTEM, UNRELIABLE SERVER,

QUEUEING SYSTEM BREAKDOWN

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

SEE REVERSE SIDE

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

332 550 JOB

Technical Report No. 188, by Andrew W. Shogan

This paper considers a single server queueing system that alternates stochastically between two states: operational and failed. When operational, the system functions as an $M/E_k/1$ queue. When the system is failed, no service takes place but customers continue to arrive according to a Poisson process; however, the arrival rate is different from that when the system is operational. Thus, both the arrival and service distributions are nonstationary. The durations of the operating and failed periods are exponential with mean $1/c\alpha$ ALPHA and Erlang with mean $1/c\beta$ BETA, respectively. Generating functions are used to derive the steady-state quantities L and W, both of which are decreasing and convex functions of c. The paper includes an analysis of several special and extreme cases and an application to a production-storage system.